CrossMark

The factor structure of a Polish language version of the hospital anxiety depression scale (HADS)

John B. Nezlek [1,2] · Marzena Rusanowska [3] · Paweł Holas [4] · Izabela Krejtz [5]

Abstract
The Hospital Anxiety Depression Scale (HADS) is widely used as a screening measure for psychological distress. Nevertheless, there is disagreement among researchers about the extent to which the HADS provides separate measures of anxiety and depression or a single measure of affectively based distress, and the present study was designed to contribute to this discussion. Participants ($n = 951$) who were psychologically distressed, but not hospitalized, completed a Polish language version of the HADS. A confirmatory factor analysis of participants' responses confirmed a model with two correlated factors with one cross-loaded item. The estimated correlation between the factors was .68. These results suggest that the Polish version of the HADS consists of two correlated measures of affectively focused distress, depression and anxiety. Nevertheless, analysts and practitioners need to be cautious and take into account the possibility that the discriminant validity of these two scales may be somewhat limited given the correlation between the subscales and the possible cross-loadings of items.

Keywords HADS · Depression · Anxiety · Factor analysis · Polish language measures

The Hospital Anxiety and Depression Scale (HADS) was developed by Zigmond and Snaith (1983) as a means of determining if an individual is experiencing psychological distress defined in terms of the two most common types of distress, depression and anxiety. The term "Hospital" in the title refers to the fact that the measure was intended to be used by hospital staff to determine if patients (originally outpatients, although this has been extended to include in-patients) are experiencing psychological distress. The scale was designed to be administered by staff who are not trained as psychiatric specialists, and the HADS was not designed to provide psychiatric diagnoses per se. The items on the scale refer to common experiences and do not use technical or diagnostic terms. The scale can be administered by a staff member, or similar to numerous other self-report measures of symptoms, it can be completed by patients/participants themselves (Zigmond and Snaith 1983).

The scale has fourteen items, seven meant to measure depression and seven meant to measure anxiety, and scores on the scale were intended to provide a quick and efficient method of determining patients' mental health in terms of these two common emotional disorders. In their review, Cosco et al. (2012) suggested a "caseness" cutpoint of 8 for both the depression and the anxiety subscales (HADS-D and HADS-A). A "caseness" cutpoint is a value that can be used to determine when an individual's level of distress might meet diagnostic criteria for a disorder, e.g., the transition from dysphoria (sub-clinical levels of depression) to depression.

The HADS has been found to have sensitivity and specificity as good as other commonly used self-rating screening instruments (Hermann 1997; Bjelland et al. 2002; Mykletun et al. 2001). Although the HADS has been widely used in populations with somatic complaints (Cosco et al. 2012), several normalization studies have been conducted in individuals with major depression symptoms (Mykletun et al. 2001; Flint and Rifat 1996; Friedman et al. 2001; Hansson et al. 2008; Matsudaira et al. 2009) and in non-clinical populations (Mykletun et al. 2001; Andrea et al. 2004).

Although the HADS is widely used, concerns have been expressed about its psychometric properties, including its

✉ John B. Nezlek
jbnezl@wm.edu

1   SWPS University of Social Sciences and Humanities, Poznań, Poland

2   College of William & Mary, Williamsburg, VA, USA

3   Polish Academy of Sciences, Warsaw, Poland

4   University of Warsaw, Warsaw, Poland

5   SWPS University of Social Sciences and Humanities, Warsaw, Poland

factor structure, item construction, and the response scales used in its administration (e.g., Coyne and van Sonderen 2012). For example, based on their meta-analysis of 75 studies that specifically examined the factorial structure of the HADS, Cosco et al. (2012) concluded that: "The heterogeneous results of the current review suggest that the latent structure of the HADS is unclear, and dependent on statistical methods invoked. While the HADS has been shown to be an effective measure of emotional distress, its inability to consistently differentiate between the constructs of anxiety and depression means that its use needs to be targeted to more general measurement of distress" (p. 180).

In contrast, based on three meta-analyses, Bjelland et al. (2002), concluded that: "HADS was found to perform well in assessing the symptom severity and caseness of anxiety disorders and depression in both somatic, psychiatric and primary care patients and in the general population" (p. 69). Although Cosco et al. (2012) acknowledged the conclusions of Bjelland et al. (2002), they thought that on balance, it is not clear if the HADS can distinguish anxiety and depression. Further clarification is needed also because most psychometric studies of the scale have been conducted on selected samples of patients with somatic illnesses (Bjelland et al. 2002; Mykletun et al. 2001), and it is less clear what the dimensional structure of the HADS is in psychiatric and non-clinically distressed populations.

The HADS has been translated into various language, and a Polish language version of the HADS was created for Polish language speakers (Karakuła et al. 1995). Polish is the official language of Poland with an estimated population in 2014 of 38,500,000 (Demographic Yearbook of Poland 2015), and there are large Polish speaking populations throughout the world. Therefore, a Polish language version of the HADS has the potential for broad application.

Although the Polish language version of the HADS has been available for some time, the psychometric properties of this measure (specifically the factor structure underlying the items) are not well understood. Previous research on Polish language versions of the HADS has produced inconsistent results (Mihalca and Pilecka 2015; Watrowski and Rohde 2014; Wichowicz and Wieczorek 2011) perhaps in part because the studies have examined very specific, limited populations (adolescents, gynecological, and post-stroke patients, respectively). Moreover, only one study (Mihalca and Pilecka 2015) used confirmatory factor analysis (CFA), and they studied only adolescent girls.

In addition to understanding the factorial structure of the Polish version of the HADS to improve our understanding of what scores on the measure represent, there is also a pressing need for a simple and easy to administer Polish language measure of anxiety and depression (psychological distress) in Polish. There is paucity of epidemiological data describing psychological distress in Poland, and one of the few epidemiological studies that has been conducted (EZOP, Kiejna et al. 2015) found

that the prevalence of major depression and other common mental disorders (CMD) in Poland was meaningfully lower than in other European countries. Major depression was estimated at only 3.0% in Poland (Kiejna et al. 2015), whereas it was estimated at 21.0% in France and 9.9% in Germany (Alonso et al. 2004), and 14.6% in Ukraine (Bromet et al. 2005).

The authors of the Polish EZOP study believed that the most likely reason for the low rates of distress in Poland was the low response rate among Polish respondents (e.g., 50.4% vs. 78.3% in Ukraine). Individuals who are suffering more intense distress are less likely to participate in a study than individuals who are suffering less intensely or who are not distressed (e.g., Graaf et al. 2000), which would lead to an underestimate of the prevalence in the population. Understanding the causes for the unexpectedly (and probably inaccurate) low prevalence of common depressive disorder in the EZOP Poland study requires additional research comparing the prevalence of distress in Poland with distress in other countries.

The HADS seems to be a good candidate for a screening tool to measure psychological distress in Poles, as it is a short and convenient self-rating test with only 14 items that provides a quick and efficient method of determining person's depression and anxiety. Before such work begins however, the factorial structure of the Polish language version of the HADS needs to be established. Does it in fact provide separate measures of depression and anxiety?

Unfortunately, there is a lack of psychometric studies evaluating the factorial structure of the Polish language version of the HADS in distressed, non-hospitalized adult population, and the primary goal of the present study was to determine if the Polish version of the HADS measured the two factors that the HADS was originally intended to measure. We collected data from a large, non-hospital sample using methods that we believed mitigated some of the shortcomings described by Coyne and van Sonderen (2012) such as problems with changing scale labels and values. Based on previous research, our general expectation was that we would find two factors, one representing symptoms of depression and another representing symptoms of anxiety. Moreover, we expected that these factors would be correlated and that some items would load on both factors.

Method

Participants in the present study were volunteers from the community who completed the HADS as part of a screening process to recruit participants for two studies.[1] Participants responded to

---

[1] One study was about cognitive processing and the other was about mindfulness training. The specific nature of these studies was not described in the call, and so we do not distinguish participants as a function of the study for which they were being screened.

offers to take part in paid studies that were posted on two Polish websites. We placed our advertisement in the section for temporary jobs and clearly stated that it was a call for paid participation in a scientific study. We were interested in studying individuals who were depressed, and so we provided information that allowed potential participants to self-screen.

The call for participants specified that we were looking for individuals who were experiencing one or more of the following symptoms (symptoms that are associated with a diagnosis of depression): 1. Were recently feeling down, sad, empty, 2. Had been feeling that everyday activities were not as pleasurable as before, 3. Were having sleeping problems, 4. Were having problems with appetite, or 5. Were feeling low in self-esteem. Participants did not answer questions about these problems; rather, we mentioned these symptoms so that potential participants had some idea regarding the basis on which participants would be selected for the studies we were conducting.
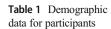
The call stated that if participants wanted to take part in the study, they would have to complete a set of questionnaires. Participants used a direct link from the advertisement to log into the website where they completed the HADS and indicated their gender, age, education, and living situation. Participants also provided information about whether they took psychiatric medications at the time of responding to the call.

Sample Demographics

The mean age of our sample of 951 was 32.1 years old ($SD = 10.6$; range 17–70), 72% (628) were women, 99% had high school degree or more, and of the 918 subjects who answered a question about current psychiatric medication 87% indicated that they were not taking any psychiatric medication. More detailed demographic information is contained in Table 1.

Measure

The Polish version of the HADS (Karakuła et al. 1995) was presented on a website. Participants were told to read each statement and to choose the response that best described how they felt during the past week. They were told that we were interested in their immediate answers to the questions and that they did not have to think too long about each answer. Responses were made using a radio button without any numbers assigned to each response option. The labels for each response option were the same as those used in the original version, although we did not present the numbers (0, 1, 2, 3) that accompanied each response option. The English language items are presented in Table 2, and the Polish language items and scales are available at https://osf.io/fwxu3/?view_only=7390795b69ce4a2c81d54b4538c463e2.

Table 1 Demographic data for participants

Educational background (950 respondents)		
Primary school only	8	0.8%
Secondary school	274	28.8%
Presently student	371	39.1%
Higher education	296	31.2%
Not specified	1	
Marital status (918 respondents):		
Married	165	18.0%
With a partner	334	36.4%
Not in a relationship	341	37.1%
Widow/widower	10	1.1%
Divorced	54	5.9%
Other	14	1.5%
Living situation (917 respondents):		
Alone	145	15.8%
With family	408	44.5%
With friends	134	14.6%
With a partner	212	23.1%
Other	18	2.0%

For each characteristic the number of individuals who selected each response in presented in the first column, and the second column contains the percent of the total number of respondents who answered this question and provided each response

Although the original HADS presented a numerical value (0, 1, 2, or 3) for each response option, we thought that such numerical values could confuse participants (e.g., Coyne and van Sonderen 2012). In the original HADS scale, for some items a higher numbered response is associated with weaker endorsement than a lower numbered response. For example, for item 1 "I feel tense or wound up" the first response is "Most of the time" and is accompanied by 3, and the last response "Not at all" is accompanied by 0. In contrast, for other items a higher numbered response is associated with weaker endorsement than a lower numbered response. For example, for item 2 "I still enjoy the things I used to enjoy" the first response is "Definitely as much" and is accompanied by 0, and the last response "Hardly at all" is accompanied by 3. We should note that all items were scored such that higher numbers indicate more distress.

We think that presenting response options without numerical values reduced the confusion experienced by respondents. This was also consistent with the recommendations of Gehlbach and Brinkworth (2011). In their discussion of best practices for designing questionnaires they explicitly recommend "Avoid using reverse-scored items" and "Because numbers have implicit meaning for many participants – which may conflict with the verbal response anchors – they should be avoided" (p. 383). It is also important to note that we did

Table 2 Univariate summary statistics for HADS items (*n* = 951)

Item –original English language	M	SD
1. I feel tense or wound up	2.03	.737
2. I still enjoy the things I used to enjoy	1.52	.766
3. I get a sort of frightened feeling as if something awful is about to happen	2.05	.813
4. I can laugh and see the funny side of things	1.34	.787
5. Worrying thoughts go through my mind	2.16	.961
6. I feel cheerful	1.25	.677
7. I can sit at ease and feel relaxed	1.82	.645
8. I feel as if I am slowed down	1.78	.819
9. I get a sort of frightened feeling like 'butterflies' in the stomach	1.52	.833
10. I have lost interest in my appearance	1.43	.925
11. I feel restless as if I have to be on the move	2.09	.831
12. I look forward with enjoyment to things	1.88	.998
13. I get sudden feelings of panic	1.28	.920
14. I can enjoy a good book or TV program	1.24	.905

not change the wording of any of the items or labels; rather, we simply did not present any numbers with the scale labels.

Results

Overview of Analyses

The analyses were done in three steps. First, univariate statistics for the items and scores for depression and anxiety subscales were calculated. These subscales were defined in terms of how the items were designed to be used. Next, a series of confirmatory factor analyses (CFA) were done to test the hypothesis that the items of the HADS in fact measure two correlated latent constructs, one representing symptoms of depression and another representing symptoms of anxiety. Finally, a series of exploratory factor analyses (EFA) were conducted to allow more direct comparisons of the present results to previous results.

Univariate Analyses

Consistent with previous research all items were coded so that higher numbers represented more distress or reduced well-being compared to lower numbers. Summary statistics and correlation matrices for the items are presented in Tables 2 and 3. The raw data are available at https://osf.io/fwxu3/?view_only=7390795b69ce4a2c81d54b4538c463e2.

We should note that traditionally scores on the HADS are calculated by summing responses to the items on each subscale. Following this procedure produced a mean anxiety score of 12.95 (*SD* = 3.79, α = .81) and a mean depression score of 10.43 (*SD* = 3.97, α = .80). The alphas for the two

subscales are similar to those reported in previous studies (e.g., Bjelland et al. 2002). The correlation between these scores was .62. These mean scores were above the values usually used to indicate that someone is experiencing levels of distress that might be clinically meaningful, which is usually a score above 8 (Bjelland et al. 2002). Using the scales as originally designed, 87% of the sample had scores of 9 or above on the anxiety subscale, and 70% had scores of 9 or above on the depression subscale. This is consistent with the self-screening of participants.

Confirmatory Factor Analyses

We conducted our factor analyses using MPlus (Muthén & Muthén 2017). Our first analysis was a confirmatory factor analysis in which we modeled two correlated factors, one consisting of the depression items (2, 4, 6, 8, 10, 12, and 14) and the other consisting of the anxiety items (1, 3, 5, 7, 9, 11, 13). The estimation method was maximum likelihood. To obtain a final model, we followed recommendations offered by Marsh et al. (2014) who suggested that the underlying structure of a set of items is best discovered by a blend of "pure" CFA (i.e., all parameters are specified in advance) and exploratory factor analyses (EFA). Marsh et al. specifically discussed the weakness of CFAs that do not model cross-loaded items when such cross-loadings may lead to improved model fits (pp. 104–15).

The initial model (with no cross loadings) fit the model reasonably well (e.g., $\chi^2(76) = 491.9$, $p < .001$, RMSEA = .08, TLI = .88), although this fit fell short of commonly accepted standards for good fits. For this model, the estimated correlation between the factors was .75. Following the modification indices suggested by the program we included a cross-loading for item

Table 3 Correlation matrix for HADS items

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		.349	.466	.298	.550	.251	.388	.332	.351	.202	.376	.308	.357	.245
2	.349		.266	.532	.300	.413	.457	.369	.260	.309	.348	.467	.248	.453
3	.466	.266		.270	.526	.217	.281	.280	.343	.168	.331	.281	.373	.256
4	.298	.532	.270		.290	.413	.460	.281	.243	.331	.318	.432	.238	.461
5	.550	.300	.526	.290		.261	.386	.328	.393	.206	.441	.360	.350	.249
6	.251	.413	.217	.413	.261		.428	.319	.197	.320	.268	.359	.320	.366
7	.388	.457	.281	.460	.386	.428		.301	.331	.294	.314	.447	.285	.420
8	.332	.369	.280	.281	.328	.319	.301		.224	.337	.399	.298	.247	.275
9	.351	.260	.343	.243	.393	.197	.331	.224		.138	.278	.299	.485	.252
10	.202	.309	.168	.331	.206	.320	.294	.337	.138		.280	.284	.198	.327
11	.376	.348	.331	.318	.441	.268	.314	.399	.278	.280		.361	.268	.254
12	.308	.467	.281	.432	.360	.359	.447	.298	.299	.284	.361		.315	.367
13	.357	.248	.373	.238	.350	.320	.285	.247	.485	.198	.268	.315		.239
14	.245	.453	.256	.461	.249	.366	.420	.275	.252	.327	.254	.367	.239	

7 (an anxiety item: "I can sit at ease and feel relaxed") on the depression factor. This model produced a χ^2 fit of 345.7 (df = 75, p < .001), a RMSEA of .062 (90% CI: .055 to .068), and a CFI of .93. We modeled this cross-loading because we believed that it led to a substantial improvement in the model fit, i.e., a difference in the χ^2 fit of 146 (df = 1, p < .001).
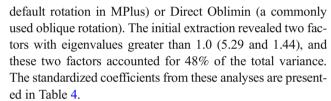
The modification indices also suggested that three other cross-loadings (items 8 and 12 on the anxiety factor and item 11 on the depression factor), would improve the model fit. Including these cross-loadings produced a model fit, a χ^2 fit of 269.5 (df = 75, p < .001), a RMSEA of .05 (90% CI: .05 to .06), and a CFI of .95. The estimated correlation between the factors in this model was .60. Nevertheless, we were not confident that the addition of these cross-loadings was justified despite the fact that the change in the fit was significant, so we adopted a final model with only one cross-loading. The standardized factor loadings and the estimated correlation between the factors are presented in Table 4.

Exploratory Factor Analyses

Much of the research on the factor structure of the HADS has relied on EFA rather than CFA. Given this, we conducted a series of EFA to provide a basis to compare more directly the present results to the results of previous research. We conducted these analyses following the guidelines offered by Costello and Osborne (2005). To reduce the extent to which our conclusions depended on the results of a specific type of analysis, we used two extraction methods, Maximum Likelihood Factor Analysis (MLFA) and Unweighted Least Squares (ULS). We chose ULS because it has been suggested that it is more robust to non-normal data (Osborne 2014, p. 10). Each of these solutions was rotated using Geomin (the oblique

default rotation in MPlus) or Direct Oblimin (a commonly used oblique rotation). The initial extraction revealed two factors with eigenvalues greater than 1.0 (5.29 and 1.44), and these two factors accounted for 48% of the total variance. The standardized coefficients from these analyses are presented in Table 4.

As can be seen from the results of these analyses, the four combinations of extraction and rotation produced very similar factor loadings, and the estimated correlations between the factors were similar (near .6). As intended, all of the odd numbered items loaded on the same factor (anxiety), and all the even numbered items loaded on another factor (depression). Nevertheless, four items cross-loaded on both factors across all four analyses. Not surprisingly, these cross-loaded items were the same items that were indicated as cross-loaded items by the modification indices in the CFA.

Additional Analyses

Although not the focus of any specific hypothesis, we examined relationships between HADS subscale scores and age, and we examined sex differences in these scores. For these analyses we used scale scores based on the raw data to provide the broadest basis for allowing researchers to compare our results to the results of other studies. We should note that scores based on the raw data were highly correlated with the factor scores from all the analyses. For example, the correlation between factor scores based on the CFA and raw scores was .99 for the Anxiety subscale and .97 for the Depression subscale.

These analyses found that women had higher scores than men on the HADS Anxiety subscale (Ms = 13.3 vs. 12.2, t(948) = 4.12 p < .001), whereas there were no sex differences for scores on the HADS Depression subscale (t < 1). Although

Table 4 Factor coefficients for items: Polish version of the HADS as a function of extraction and rotation

	Factor extraction and rotation									
	CFA		MLFA Geomin		MLFA Oblimin		ULS Geomin		ULS Oblimin	
Item	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Anxiety										
1A	.70		.66		.67		.65		.66	
3A	.65		.67		.69		.67		.69	
5A	.75		.78		.80		.77		.78	
7A	.14	.57	.17	.56	.16	.56	.16	.56	.15	.56
9A	,56		.49		.50		.53		.54	
11A	.57		.40	.25	.40	.23	.39	.25	.39	.25
13A	.55		.46		.46		.50		.50	
Depression										
2D		.72		.72		.73		.71		.73
4D		.69		.73		.74		.72		.74
6D		.60		.59		.60		.60		.61
8D		.50	.25	.34	.25	.33	.24	.35	.24	.35
10D		.48		.48		.49		.50		.51
12D		.64	.16	.53		.54	.16	.52		.53
14D		.52		.64		.66		.64		.68
F1*F2	.68		.57		.61		.58		.61	

For EFA, factor coefficients of .15 or less are not displayed

F1 first factor, *F2* second factor, *A* anxiety item, *D* depression item, *F1*F2* estimated correlation between the factors

we also had no hypotheses about correlations between HADS scores and age, we found that participant age was not significantly correlated with scores on the HADS Anxiety subscale ($r = .04$), whereas age was positively correlated with scores on the HADS Depression subscale ($r (948) = .14$, $p < .001$).

Discussion

We found that the Polish language version of the HADS when examined using an adult, psychologically distressed, but non-hospitalized population, had two (correlated) factors that corresponded reasonably closely to the intended dimensions of anxiety and depression. The current findings of a model with two correlated factors and an estimated correlation between the factors of .65 corroborates findings that examined similar samples, i.e., psychological distressed members of the community (e.g., Mykletun et al. 2001; Matsudaira et al. 2009). Moreover, the cross-loading in our CFA of an item on the depression factor that was meant to measure anxiety (7: "I can sit at ease and feel relaxed") is consistent with the results of numerous previous studies (e.g., Cosco et al. 2012). In fact, in their review Cosco et al. singled out this item as one of the most troublesome items on the scale (p. 182).

The present results are at odds however with some research of psychiatric populations such as Friedman et al. (2001)

which suggests that the HADS has three factors, two reflecting anxiety ('psychic anxiety', 'psychomotor agitation') and one reflecting depression. We should note that although our sample was psychologically distressed, they did not receive a formal psychiatric evaluation, nor were they hospitalized, allowing for the possibility that differences between the present results and the results of Friedman et al. (2001) may be due to differences in the samples examined in the two studies. It should be noted that Friedman et al. conducted EFAs with two factors and found factor loadings that were consistent with the present two-factor model (p. 251).

Although some previous research on Polish language samples has found three factors (Mihalca and Pilecka 2015; Watrowski and Rohde 2014), it is not clear how to interpret these results. For example, the third factor reported by Watrowski and Rohde (2014) had only two items with loadings greater than .5. Moreover, they presented only the results of a principal component analysis (PCA) with an orthogonal rotation. An oblique rotation may have provided a meaningfully different set of coefficients. Moreover, despite finding three factors, they concluded that the HADS was probably best used as "as an indicator for global psychological distress" (p. 517). Although Mihalca and Pilecka (2015) used CFA to examine the factorial structure, they studied only adolescent girls. Regardless, we did not find any compelling evidence for three factors. The eigenvalue of the third factor

from our EFA was below the commonly accepted cutpoint of 1.0, and the fit indices of the CFA did not suggest the existence of a third factor.

There is an active debate about whether the HADS is best characterized as a multifactorial measure or is best thought of a measure of general distress. For example, based on their review of 50 studies, Cosco et al. (2012) concluded that" While the HADS has been shown to be an effective measure of emotional distress, its inability to consistently differentiate between the constructs of anxiety and depression means that its use needs to be targeted to more general measurement of distress" (p. 180). It is important to note that this conclusion of Cosco et al. was not based on the number of factors found in various studies; quite the opposite, they reported only 1 study that reported 1 factor. Rather, their conclusion was based on the inconsistency in the numbers of factors that have been found (sometimes two, sometimes three) and on the inconsistency in the factor structures that have been found (numerous items cross-load on factors and do so inconsistently). Cosco et al. (2012) do not argue with the conclusion that the HADS measures distress. They simply did not think that the HADS is precise enough to provide discriminatively valid measures of depression and anxiety (Cosco et al. 2012; p. 184).

Although we appreciate the point Cosco et al. and others have made about the precision of the HADS, at this point, taking the available research into account, the present study suggests that the Polish language version of the HADS measures two factors, anxiety and depression, with same caveats regarding cross-loadings of some items and the correlation between the two factors. Sex differences in scale scores and the correlations between age and scales scores provided additional support for the conclusion that the HADS measures two constructs. Women were more anxious then men, but there were no differences in depression, and age was correlated with depression scores but not with anxiety.

If the subscales of the HADS are virtually interchangeable manifestations of a general construct of distress then correlations between the two subscales and other measures should be similar and group differences that occur for one subscale should occur for the other. This did not occur in the present study despite the fact that there was adequate power to detect a correlation between age and anxiety and a difference between men and women in terms of depression. Using G-power (Faul et al. 2009), the estimated power of finding a correlation of .15 was .99, and it was .98 for finding a mean difference between men and women the size that was found for anxiety.

There is also the issue of the correlation between the factors, approximately .6 across the most of the analyses we conducted. Some could argue that such a correlation indicates that the factors are simply manifestations of an underling factor of general distress. Nevertheless, even if one assumes that the correlation is .7 (the estimated correlation between the latent factors in our CFA was .68, and it was corrected for error), this

would mean that the two factors share approximately 50% of the variance. Most important, this also means that they do not share 50%.

Two measures that have 50% of their variance in common clearly measure related constructs; nevertheless, there is ample unshared variance to provide a basis for claiming that the two measures measure constructs that can be distinguished. Anxiety and depression are co-morbid; yet affectively speaking, depression involves deactive negative affect, whereas anxiety involves active negative affect. There is overlap between the two (both involve negative affect), but they are also distinct in terms of level of arousal.

In terms of the factor loadings, our initial CFA found a reasonable fit (RMSEA = .076, CFI = .90) with no cross-loaded items and no correlated error terms. We added a cross-loading for item 7 to improve the RMSEA to reach the .05 level, a commonly accepted minimum for model fit. This means that the covariances among the HADS items can be interpreted as manifestations of two factors, as intended. Unfortunately, demonstrating that the HADS measures two factors is not enough to validate the measure. We did not have the data necessary to determine the relationships between HADS scores and formal diagnosis, which are necessary to confirm the validity of the measure. This is something that future research needs to address. Nevertheless, demonstrating that scores on the Polish language version of the HADS represent distinguishable constructs is a necessary pre-condition for establishing its validity as a measure of depression and anxiety.

Conclusion

The current results suggest that the HADS distinguishes anxiety and depression, and by extension, the present results suggest that the Polish language version of the HADS may serve as a convenient screening tool to measure psychological distress in Poles. Short, easy to complete self-rating scales in Polish are needed. The only cross-sectional epidemiological study of mental problems in Poland of which we are aware (EZOP, Kiejna et al. 2015) described implausibly low rates of the prevalence of depression and anxiety disorders among Poles, particularly in comparison other European countries. Such findings call for additional assessment of the prevalence of depression and anxiety in Poland, and we believe that HADS can be a good candidate for the preliminary selection of psychologically distress individuals and who might meet diagnostic criteria for a disorder.

The present results are limited to a non-hospitalized adult population that was selected on the basis of presently experiencing psychological distress, and so the present results may not be generalizable to other Polish residents who are not experiencing psychological distress or who suffering from a general medical condition. Moreover, the convergent validity

of the Polish version of the HADS should be confirmed by examining its relationships to other anxiety and depression scales. For example, the HADS focuses on symptoms of autonomic anxiety and anhedonic depression, whereas other scales (e.g., Beck Depression Inventory and State-Trait Anxiety Inventory) measure other aspects of depression and anxiety such as guilt, helplessness, and somatic symptoms. The HADS explicitly excludes somatic symptoms. Nevertheless, consistent with the recommendation of other researchers (e.g. Bjelland et al. 2002), our results can be interpreted as providing support for considering the HADS as originally intended, i.e., as an instrument that provides distinct measures of depression and anxiety; however the correlation between these two subscales and the possibility that some items load on both factors need to be kept in mind if the scales are used in this way.

Compliance with Ethical Standards

Ethics Review This study was approved by the Institutional Review Board for research involving human subjects of SWPS University of Social Sciences and Humanities.

Informed Consent Informed consent was obtained from all participants in this study.

Conflict of Interest The authors declare that they have no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H., et al. (2004). Investigators, European study of the epidemiology of mental disorders (ESEMeD) project. Prevalence of mental disorders in Europe: Results from the European study of the epidemiology of mental disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica. Supplementum, 109*, 21–27. https://doi.org/10.1111/j.1600-0047.2004.00327.x.

Andrea, H., Bültmann, U., Beurskens, A. J., Swaen, G. M., van Schayck, C. P., & Kant, I. J. (2004). Anxiety and depression in the working population using the HAD scale-psychometrics, prevalence and relationships with psychosocial work characteristics. *Social Psychiatry and Psychiatric Epidemiology, 39*, 637–646.

Bjelland, I., Dahl, A., Haug, T., & Neckelmann, D. (2002). The validity of the hospital anxiety and depression scale: An updated literature review. *Journal of Psychosomatic Research, 52*, 69–77.

Bromet, E. J., Gluzman, S. F., Paniotto, V. I., Webb, C. P., Tintle, N. L., Zakhozha, V., et al. (2005). Epidemiology of psychiatric and alcohol disorders in Ukraine. *Social Psychiatry and Psychiatric Epidemiology, 40*, 681–690.

Cosco, T. D., Doyle, F., Ward, M., & McGee, H. (2012). Latent structure of the hospital anxiety and depression scale: A 10-year systematic review. *Journal of Psychosomatic Research, 72*, 180–184. https://doi.org/10.1016/j.jpsychores.2011.06.008.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10(7)*. Available online: http://pareonline.net/getvn.asp?v=10&n=7

Coyne, J. C., & van Sonderen, E. (2012). No further research needed: Abandoning the hospital and anxiety depression scale (HADS). *Journal of Psychosomatic Research, 72*, 173–174. https://doi.org/10.1016/j.jpsychores.2011.12.003.

Demographic Yearbook of Poland (2015). Central Statistical Office of Poland. http://stat.gov.pl/en/topics/statistical-yearbooks/statistical-yearbooks/demographic-yearbook-of-poland-2015,3,9.html. Accessed 07 Oct 2015.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.

Flint, A. J., & Rifat, S. L. (1996). Validation of the hospital anxiety and depression scale as a measure of severity of geriatric depression. *International Journal of Geriatric Psychiatry, 11*, 991–994.

Friedman, S., Samuelian, J. C., Lancrenon, S., Even, C., & Chiarelli, P. (2001). Three dimensional structure of the hospital anxiety and depression scale in a large French primary care population suffering from major depression. *Psychiatry Research, 104*, 247–257.

Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology, 15*, 380–387. https://doi.org/10.1037/a0025704.

Graaf, R. D., Bijl, R. V., Smit, F., Ravelli, A., & Vollebergh, W. A. (2000). Psychiatric and sociodemographic predictors of attrition in a longitudinal study the Netherlands mental health survey and incidence study (NEMESIS). *American Journal of Epidemiology, 152*, 1039–1047.

Hansson, M., Bodlund, O., & Chotai, J. (2008). Patient education and group counseling to improve the treatment of depression in primary care: A randomized controlled trial. *Journal of Affective Disorders, 105*, 235–240.

Hermann, C. (1997). International experiences with the hospital anxiety and depression scale-a review of validation data and clinical results. *Journal of Psychosomatic Research, 42*, 17–41.

Karakuła, H., Grzywa, A., Spiła, B., & Jedrych, M. (1995). Zastosowanie Skali Lęku i Depresji w chorobach psychosomatycznych. [Use of Hospital Anxiety and Depression Scale in psychosomatic disorder.]. *Psychiatria Polska, 30*, 653–668.

Kiejna, A., Piotrowski, P., Adamowicz, T., Moskalewicz, J., Wciórka, J., Stokwiszewski, J., et al. (2015). The prevalence of common mental disorders in the population of adult Poles by sex and age structure–an EZOP Poland study. *Psychiatria Polska, 30*, 15–27.

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85–110.

Matsudaira, T., Igarashi, H., Kikuchi, H., Kano, R., Mitoma, H., Ohuchi, K., & Kitamura, T. (2009). Factor structure of the hospital anxiety and depression scale in Japanese psychiatric outpatient and student

populations. *Health and Quality of Life Outcomes, 7*, 42. https://doi.org/10.1186/1477-7525-7-42.

Mihalca, A. M., & Pilecka, W. (2015). The factorial structure and validity of the hospital anxiety and depression scale (HADS) in Polish adolescents. *Psychiatria Polska, 49*, 1071–1088. https://doi.org/10.12740/PP/38139.

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (8th ed.). Los Angeles: Muthén & Muthén.

Mykletun, A., Stordal, E., & Dahl, A. A. (2001). Hospital anxiety and depression (HAD) scale: Factor structure, item analyses and internal consistency in a large population. *British Journal of Psychiatry, 179*, 540–544.

Osborne, J. W. (2014). *Best practices in exploratory factor analysis*. Scotts Valley: CreateSpace Independent Publishing.

Watrowski, R., & Rohde, A. (2014). Validation of the polish version of the hospital anxiety and depression scale in three populations of gynecologic patients. *Archives of Medical Science, 10*, 517–524. https://doi.org/10.5114/aoms.2013.36520.

Wichowicz, H., & Wieczorek, D. (2011). Badanie przesiewowe depresji poudarowej z użyciem Hospital Anxiety and Depression Scale (HADS). [Screening post-stroke depression using the Hospital Anxiety and Depression Scale.]. *Psychiatria Polska, 44*, 505–514.

Zigmond, A., & Snaith, R. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*, 361–370.